Miguel, Amblard, Barceló & Madella (eds.) Advances in Computational Social Science and Social Simulation
Barcelona: Autònoma University of Barcelona, 2014, DDD repository <http://ddd.uab.cat/record/125597>

Social Contagion and Homophily within Romantic Network

Francisco Linares
Departamento de Sociología,
Facultad de CC. PP y Sociales
38071 La Laguna, SPAIN
Email: flinares@ull.es

□

***Abstract*—This paper presents an agent based simulation model which attempts to show how the diffusion of a cultural-trait can be affected by an uneven distribution of influence-capacity among individuals. For the sake of concretion the model represents a population of teenagers who attempt to find a romantic relationship looking for a partner within their friendship ties. Partner choice is ruled by a homophilic principle (agents look for someone who is similar to them in a given socio-cultural trait, given a certain range of tolerance to difference). It is shown how the assignment of an especial weight on partner's influence can affect the contagion process, even if the number of romantic relationships represents a small fraction of the total number of links.**

I. Introduction

THE aim of this paper is to present an agent based simulation model, which attempts to show how the diffusion of a cultural-trait can be affected by an uneven distribution of influence-capacity among individuals. In addressing this problem two theoretical concepts are central: contagion and homophily.

"Contagion" is a fairly well known social phenomenon. Since Coleman, Menzel and Katz celebrated study on the diffusion of the use of "gammanym" among doctors [1], it is widely accepted that the influence of peers on individuals' decision to accept or refuse a given socio-cultural trait produces a kind of "snow-ball process" that usually can be represented with a typical S-shaped diffusion curve, being the speed of the process dependent on certain characteristics such as the characteristic of the "critical mass" or the network topology. This process usually ends with a fairly large proportion of the population adopting the new trait [2].

Homophily is a basic principle of structuration of social relations which means that similar individuals are connected among themselves more often than dissimilar ones, a well documented pattern in many realms of social life [3, 4]. This tendency may be the product of the distribution of population over relevant social attributes [5], the structuring impact of social foci of interaction on individuals' networks [6], or the preferences of individuals for similar others [7]. Whatever its cause, homophilic patterns have the implication of larger homogeneity in social relations than would otherwise be expected.

For the sake of concretion the model presented in this paper represents a population of teenagers who attempt to find a romantic relationship looking for a partner within their friendship ties. Partner choice is ruled by a homophilic principle which is, in the case of this artificial society, assumed to operate in a very simple way: agents look for someone who is similar to them in a given socio-cultural trait, given a certain range of tolerance to difference. At the same time, the value of this trait (which is assumed to be measured in a quantitative scale) for every agent is influenced by the values of other agents' traits in its immediate environment but –and this is the central issue of the paper- if the agent is engaged the influence of its partner is assumed to have an especial weight (which is also assumed to be measured in a quantitative scale). The statistical analysis of the model's outputs show that this uneven influence indeed has an effect on the contagious effect. What makes this result noteworthy is the fact that, because the network structure is assumed to be fixed and the rules of making romantic relationship are so restrictive (see the description of the model below), it emerges as the product of the behavior of a very small fraction of the population.

The paper will proceed as follows: First, the computational model is briefly described. Second, main results of the analysis of the model are presented. Then a discussion of the results follows. Finally, the paper ends with a short section of conclusions.

II. A Social Contagion and Partner Choice Model

In order to analyze how the dynamics of homophily and contagion interact, I have built an agent based simulation model (which I call SCPCM) where diffusion of a trait and

□ Financial support is acknowledged from research projects CSO20012-31401 (MINECO) and CSD2010-0034 CONSOLIDER-INGENIO (MICINN).

partner choice evolve at the same time within a population of 200 agents embedded in a network of 20 links per node on average. The structure of the network is kept constant (that its, there is no network evolution) in order to not to confound the effects of variation of agent's behavior and the possible effect induced by the variation of network topology. The model is fully described in the appendix, following the ODD protocol designed by Railsback and Grimm [8]. A brief description of SCPCM is provided in this section. Afterwards some hypothesis are suggested.

A. Brief description of SCPCM.

The program was implemented in the platform Netlogo [9], and reproduces the following steps (see Diagram 1 below):

1. The social network is seeded
2. One of the agents is randomly chosen.
3. If the agent has not a partner, it is asked to look for someone according to the following rules:
   - The partner must be found among link neighbors
   - The partner must be of a different sex.
   - If the agent' sex is male the partner must be younger; and the other way round if it is female.
   - The difference between the values of the trait must be within a range of tolerance which is set by a tolerance-parameter (τ).
4. If a partner is found, both agents engage. This relationship may be broken with a probability which is set by a breaking-probability-parameter (β).
5. Whether a partner is found or not, the agent is influenced by their link neighbors according to the following rules:
   - If the agent is not engaged, the value of its trait becomes the median of its link neighbors.
   - If he agent is engaged, the value of its trait is determined by both the value of the trait of its partner, weighted by a weight-parameter (ω), and the median of its link neighbors, weighted by 1 - ω .
6. Return to 2 until the process is reiterated 1,100 times.

In summary, the model contains two different mechanisms of social interaction: On the one hand, agents select their partners following a homophilic rule. The homophilic strength of the choice is determined by the parameter τ, which ranges from 0 to 1. On the second hand, agents are influenced by other agents they are tied to, so the values of their traits converge to a central value of the local environment. This contagious process is, nevertheless, affected by previous partner selection, since the value of trait of partner has a special weight. The strength of partner's influence, relative to other agents' influence, is determined by parameter ω, which also ranges from 0 to 1.

In the end there is a feed-back process between partner choice and trait contagion: the distribution of trait values influences agent's partners' pool; and, at the same time, agents' choices of partner influence the distribution of trait values. The model attempts to show the outcomes of these reinforcing flows, paying special attention to the fact that the variation among agents' values of trait within the network is determined by parameters τ, β and ω.

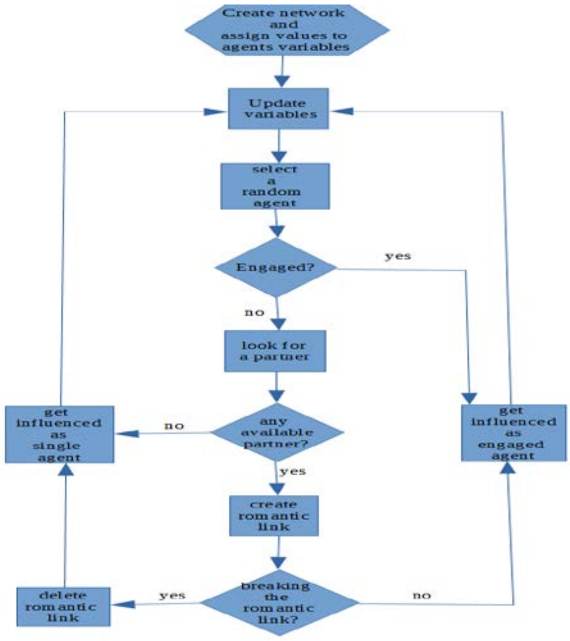


Diagram 1

B. Hypothesis.

Concerning the contagious process, variation in the distribution of trait should be positively associated with tolerance, since tolerant individuals will be "comfortable" in a world with high diversity. It should also be negatively associated with weight of partners influence, since if may partner has a strong influence on my, overall diversity is reduced. Nevertheless, there is not an obvious way to relate it to the probability of breaking a relationship. Therefore it can be expected that:

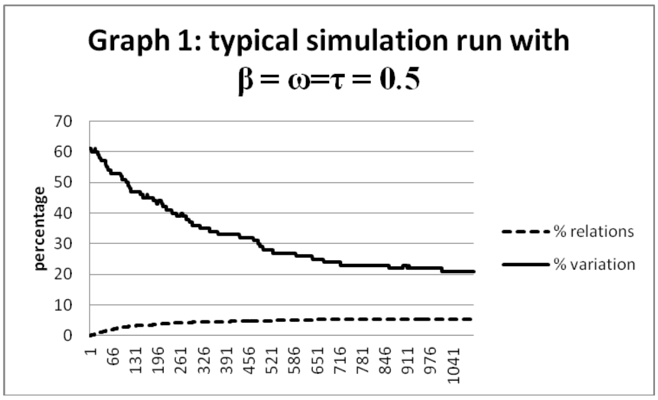
- The higher τ, the higher coefficient of variation of the trait (H1).
- The coefficient of variation of the trait will not be sensitive to β (H2).
- The higher ω, the lower the coefficient of variation of the trait (H3).

On the other hand, concerning the partner choice process, it is straight forward that as the probability of breaking romantic relationships increases, the final number of relationships must also increase. It would also seem quite obvious that the higher the tolerance to trait's difference the number of relationships should also increase. Therefore it can be expected that:

- The higher τ, the higher the number of relations (H4).
- The higher β, the higher the number of relations (H5).
- The number of relations will not be sensitive to ω (H6).

III. RESULTS

A series of simulation experiments exploring the parameter space of τ, ω and β were conducted; reiterating the simulation 50 times for every experimental condition, which amounts to 66,500 simulation runs. Results of these experiments show the emergence of patterns which are quite different from the simpler models, where every dynamic operates independently. Graph 1 displays the evolution of coefficient of variation of trait (%) and of number of relations (as a percentage of total friendship links) in a typical simulation run. The trends are quite clear: the variation of trait continuously decreases from roughly 60% to roughly 20% as the simulation progresses; meanwhile a number of romantic relationships emerge on the early stages of the simulation, and although some of them will disappear and new ones will appear the rate to total friendship relations will be kept mainly constant throughout the simulation run at a value of roughly 5%.

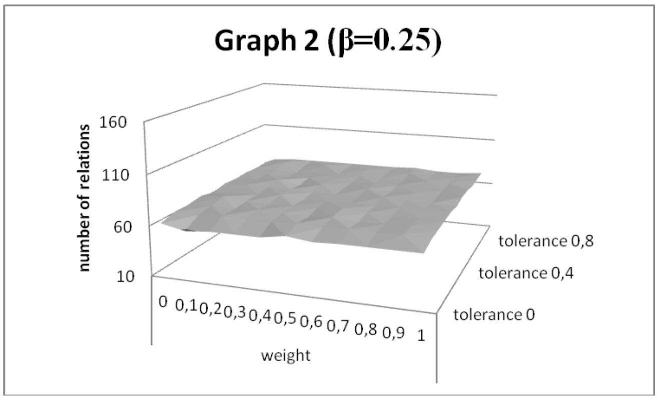


Graph 1: typical simulation run with β = ω=τ = 0.5

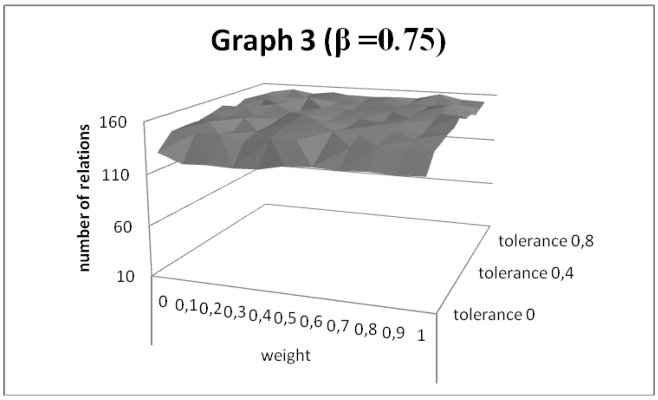
A. Number or romantic relationships.

Concerning the number of romantic relationships, the simulation provides clear support for hypothesis H5 and H6, as can be easily shown in graphs 2 and 3, which represent the number of final relationships for every combination of the spectrum parameter of τ and ω, when β equals 0.25 and 0.75 respectively[1]. It is quite obvious that variation in parameter β has the expected effect: the higher the probability of breaking a relationship, the higher the number of final relationships. It is not only the expected effect but also the greatest effect, since parameters ω and τ do not seem to have any influence. This result is clearly counter-intuitive, since one would expect the number of relations to increase with tolerance to trait of partner, as suggested by H4.

[1] Results shown in graphs are the mean values of 50 repetitions for every combination of parameters.



Graph 2 (β=0.25)



Graph 3 (β =0.75)

The linear multivariable regression model estimated for this dependent variable clearly confirms the impression produced by the graphs. "Probability of breaking a relationship" has the strongest significant effect on the dependent variable, while "weight" has no significant effect at all and "tolerance" has a very weak (although significant) effect, as shown in Table 5.1, which displays the results of the model. The value of R square for the model 0.547.

TABLE 1. DEPENDENT VARIABLE: NUMBER OF ROMANTIC RELATIONS

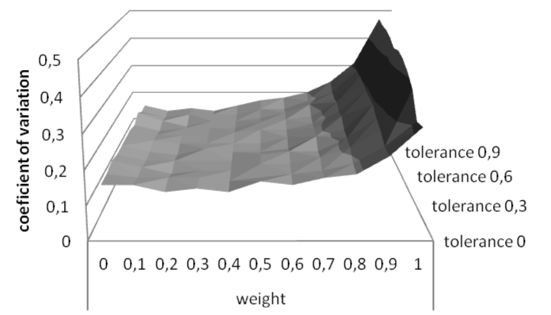
Model	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1 (Constant)	30.075	.171		176.289	.000		
probability-of-breaking-up	109.753	.203	.738	540.454	.000	1.000	1.000
tolerance	8.801	.185	.065	47.606	.000	1.000	1.000
weight	-.248	.185	-.002	-1.338	.181	1.000	1.000

B. Coefficient of variation of trait.

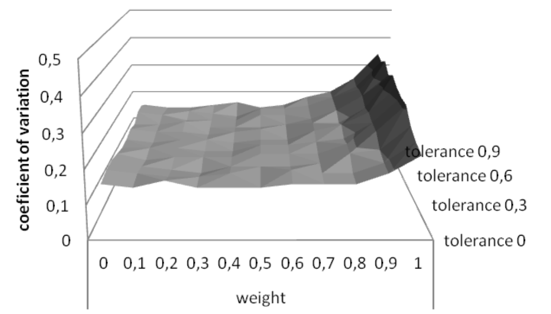
The results concerning the coefficient of variation of trait are more counter-intuitive. The variation of trait among agents increases as β decreases, but only with high values of both ω and τ. If weight of partner's trait is high, but tolerance to partner's difference is low or, the other way round, tolerance is high but weight is low, the probability of breaking a relationship does not seem to have an effect on the coefficient of variation.

Tolerance to partner's trait and weight of partner's influence have very different effects. On the one hand the coefficient of variation does not seem to be very sensitivity to the values of parameter τ. On the other hand, parameter ω seems to have the strongest influence, as coefficient of variation of trait clearly increases the higher the values of ω.

Graph 4 (β=0.25)



Graph 5 (β=0.75)



The linear multivariable regression model estimated for this dependent variable shows (see Table 2) that "weight" has the strongest significant effect on the dependent variable, but it is negative rather than positive. The effect of "tolerance" has the expected direction, although it is a rather small. Contrary to our expectations "probability of breaking a relationship" has a significant, even if small, negative effect. The value of R square for the model is just 0.192, implying that the model poorly captures the logic behind the variation of the dependent variable.

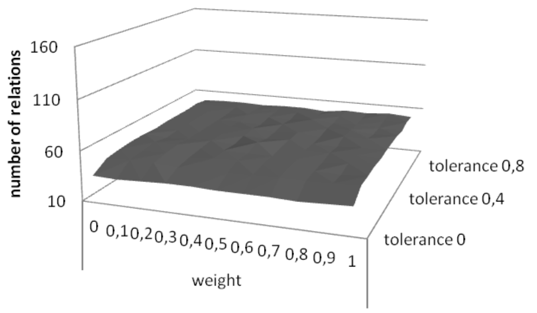
TABLE 2. DEPENDENT VARIABLE: COEFICIENT OF VARIATION OF TRAIT

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
1 (Constant)	.181	.001		353.862	.000		
probability-of-breaking-up	-.030	.001	-.090	-49.124	.000	1.000	1.000
tolerance	.018	.001	.059	32.170	.000	1.000	1.000
weight	.129	.001	.424	232.622	.000	1.000	1.000

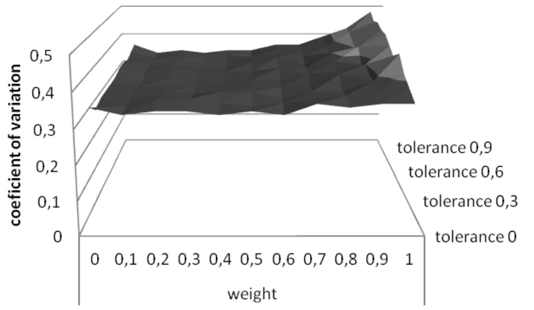
C. Sensitivity to average number of ties.

A straightforward question over these results is whether they are dependent (and if so, to what extent) on the topology of the network. (As explained in the appendix, the network is created by means of an algorithm which randomly assigns links to agents until the number of links per agent fits a certain average node degree parameter, which throughout the whole range of simulations has been set to 20).

Graph 6 (average node degree = 4)



Graph 7 (average node degree = 4)



In order to answer this question, new simulations were conducted varying the average node degree. Graphs 6 and 7 show the number of relations and coefficient of variation when the average node degree equals 4 (i.e. agents have 4 links on average) and the probability of breaking a relationship equals 0.5. The influence of average node degree on the number of relation is obviously a deterministic outcome of the model: since agents choose their partners among their linked neighbors, the less the number of ties the less the number of romantic relationships. The estimated regression model (see Table 3) shows a strong significant effect of this variable. R square for this model raises up to 0.743 .

TABLE 3. DEPENDENT VARIABLE: NUMBER OF ROMANTIC RELATIONS

Model	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1 (Constant)	-12.346	.162		-76.092	.000		
probability-of-breaking-up	109.790	.153	.738	716.837	.000	1.000	1.000
tolerance	8.773	.139	.065	62.921	.000	1.000	1.000
weight	-.253	.140	-.002	-1.811	.070	1.000	1.000
average-node-degree	3.394	.008	.442	429.131	.000	1.000	1.000

On the other hand, the influence of average node degree on the coefficient of variation is less straightforward. The linear multivariable regression model estimated for the coefficient of variation (see Table 4) shows that “average node degree” has the strongest significant effect, which is negative. The effects of “weight”, “probability of breaking a relationship” and “tolerance” are similar to the model shown in Table 5.2 above. The value of R square raises in this model up to 0.507.

TABLE 4. DEPENDENT VARIABLE: COEFICIENT OF VARIATION OF TRAIT

Model	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1 (Constant)	.302	.001		599.056	.000		
probability-of-breaking-up	-.030	.000	-.090	-63.098	.000	1.000	1.000
tolerance	.018	.000	.059	41.360	.000	1.000	1.000
weight	.129	.000	.424	297.775	.000	1.000	1.000
average-node-degree	-.010	.000	-.561	-393.762	.000	1.000	1.000

IV. DISCUSSION

Summarizing, in SCPCM the probability of breaking a relationship has a very strong positive effect on the number of relations, and a weak (but still significant effect) negative effect on the variation of trait. Tolerance to others has a positive significant effect on both variables, but in every case this effect is rather small. The strongest effect on variation of trait comes from the average node degree and the weight of partner influence (which has no effect at all on the number of relations).

There are a number of counter-intuitive results that should be stressed. One would expect higher levels of trait variation the higher the tolerance; however tolerance to others has no strong effect on trait variation. In the model the coefficient of variation invariably falls, mainly driven by the number of available ties per agent and the weight of partners’ influence. These effects are also counter-intuitive, insofar as one would expect the influence of these variables to operate on the opposite direction than they actually do. It should be expected the variation of trait to increase as the number of different neighbors also increases, and to decrease as the weight of partner’s influence increases. But the statistical tests show that it actually happens the other way round. Why is this so?

The dynamic of the simulation model allows us to understand this puzzle. Because the process of contagion is necessarily stronger the denser the network of ties, the diversity among agents is reduced (and, at a network level, the emergence of a number of trait-clusters is fostered). A result which has already been observed in previous models [10, 11]. Furthermore, when the influence of the (similar) partner has a higher weight than the influence of other linked neighbors, homophilic choices of partners seem to reinforce the homogenization effect of social contagion. I do not know of previous research accounting for this effect.

Finally, an important result of the analysis is, since the effect of “weight” on the variation of trait necessarily depends on the number of romantic relationships created through the simulations, and because this number is only a small proportion of the total amount of relations (unlikely to be higher than 5%), it follows that the behavior of a small number of agents have a strong impact on the evolution of the whole system, what is an usual feature of complex adaptive systems.

V. CONCLUSION

Contagion models usually relay on the assumption that an agent is “infected” if the rate of infected neighbors is above a certain individual threshold. The model presented in this paper departs from this basic scenario in two ways: Firstly, agents do not show a dichotomic trait (so they cannot be classified just as “infected” or “not infected”, but they differ on a continuous scale measuring a certain trait (e.g. a given musical taste), and secondly they are not equally sensitive to all its neighbors (i.e. one of its neighbors has an especial “weight” on agent’s decision to change its trait). Our analysis suggest that attention should be paid to “power” differences (especial weights) on the influence process driving social contagion, since it may have a significant impact on it.

APPENDIX: ODD PROTOCOL

A. Overview

1. *Purpose*: the model has the purpose of exploring how two different social dynamics, diffusion of a cultural trait and romantic matching, influence one each other. The specific problem the model addresses is: how these processes are both dependent on interaction based on three individual characteristics: sensitivity to others' similarity, influence of partner on owns decisions, and likelihood of breaking a romantic relationship. The model explores this dynamics in a fixed network of 200 agents which are intended to represent teenagers who hold friendship relationships which may evolve, if the right partner is found, to romantic relationship.

2. *Entities, state variables, and scales*: the model has three kinds of entities: boys, girls and links. The environment consists of a torus of 81x81 patches which have no state variable. All agents, whether boys or girls, have the following state variables: sex (boolean), age (numerical), engaged? (boolean), partners-memory (list), trait (numerical), and influence-threshold (numerical). Links represent the type of relationship between two agents by means of a color code (see below).
   Global variables are: number-of-romantic-relationships (numerical) and trait-variability (numerical), which are the main outputs of the model. Other global variables are set as parameters: likelihood-of breaking-a-relationship (numerical) tolerance-to-cultural-difference (numerical) and weight-of-partner-influence (numerical). All three of them are key parameters to explore in the model. Besides them, average-node-degree (numerical) and mean-influence-threshold (numerical) are parameters who control the average number of ties per agent, and the average sensitivity of agents to social influence.
   There are no temporal or spatial scales, since real time and/or real environment are not simulated.
3. *Process overview and scheduling*: The model includes the following actions executed every time-step in the same order:
   1. One agent is randomly chosen.
   2. If the agent is not engaged he is asked to look for a partner.
   3. If a partner is found the agent is asked to engage.
   4. Whether the agent is engaged or not it is always asked to be culturally influenced (i.e. change the value of its cultural trait).
   5. Variables are updated.

   The simulation stops after 1100 time steps, which is enough for the model to reach an equilibrium point.

B. Design Concepts

4. *Design Cocepts:*
- *Basic Principles*: The model attempts to capture the interaction of two different mechanisms: homophily and contagion. Homophily is the principle by which people tend to make relations with other people similar to them in certain traits. For the sake of simplification only one trait is represented. Contagious is a process which produce the spread of a certain trait among a population by means of social influence. In the model agents look for a romantic partner similar to them in a certain cultural trait, which is measured in a quantitative scale. At the same time agents are also influenced by their relationships, whether romantic or friendship, although these two different sources of influence do not have the same weight.
- *Emergence*: The model shows how the dynamics of romantic-matching and social influence are interdependent so the rate of variation of the cultral trait among the population and the number of romantic relationships both differ from the scenario where these two processes are independent.
- *Adaptation*: Agents perform two kinds of adaptive behavior. They become engaged if there is an agent in their local environments who meet the conditions to be chosen as a partner (details below). Second, agents change the value of their trait by means of a social influence process (details below).
- *Objectives*: There is not a fitness or utility measure in the model to be optimized. However agents behave as if they had the goal of finding a romantic partner.
- *Learning*: Agents do no learn from past experience.
- *Prediction*: Agents do not predict future conditions.
- *Sensing*: All agents occupy a position in a network, which is assumed no to evolve as time progresses. The network represents the web of friendship relationships among teenagers. When searching for a partner and when updating the value of its cultural trait, every agent has access to states variables of its local environment (i.e. other agents it has a direct tie with).
- *Interaction*: Boys and girls in the same local environment interact making (and breaking) romantic relationships (see details below). All agents in the same local environment interact influencing one another on the value of their cultural traits (see details below).
- *Stochasticity*: Stochastic processes are used in the initialization in different ways. The social network is seeded with random number 1111, in order not to confound the effect of variation on the network's topology with the effect of agents' behavior. State variables of agents are randomly initialized in every simulation run. The agent behaving in every simulation run is also randomly chosen. Since there are 200 agents and the simulation lasts for 1200 ticks, every agent has on average 6 chances of making a relationship and being influenced. Random numbers are also used in some submodels (see details).
- *Collectives*: There are two agent sets: boys who may engage with girls younger than them; and girls who may engage boys older than them. Both, boys and girls, are subject of social influence in the same way.
- *Observation*: At the end of every simulation run required outputs are: a) number of social relationships made up through the simulation; b) actual coefficient of variation of the cultural trait. Plots show the evolution of these indicators through time steps. Besides that, it is also shown in the interface whether a certain link represent

friendship (black links), a current romantic relationship (green links) or a past relationship (grey links). Agents are represented by means of circles whose color shows the value of the cultural trait (from light gray for low values to dark gray for high values).

C. Details

5. *Initialization*: The simulation is initialized with 200 agents, whose state variables are randomly assigned. Sex is assigned with 50% chance. Age of agents is picked up from a uniform distribution within the range 14 to 17. Trait of agents is picked up from a uniform distribution within the range 0 to 9. Influence-threshold is set by a parameter between 0 and 1 (currently set to 1, i.e maximum sensitivity to influence). The variable engaged? is set false for all agents. Memory of past partners is initially empty.

   Then, links are then created with a random seed; the random assignment of links to agents ends when the condition of 20 links per agents on average is met. This produces a small-world type of network. The procedure is copied from Stonedhal and Wilensky (2008).
6. *Input data*: no input data are required.
7. *Submodels*:

- *Look for a partner:*

If an agent is selected to look for a partner it will randomly pick, if any, one of his link neighbors which meet tall three conditions:

a) opposite sex

b) if agents is a man, partner must be younger. If it is a woman, partner must be older.

c) the absolute difference between the two values of trait divided by ten must be less than the value set by the parameter tolerance. This grants that agents will engage with agents with very similar value in trait when tolerance is low; but pool of possible partners will be larger when tolerance is high.

- *Get engaged:*

If a partner has been selected, the agent checks that it is not a member of the list of previous partners. Then it includes partner in this list, change the state of engaged? to true, and asks partner to do both actions. However if a random number extracted from a uniform distribution between 0 and 1 is below the value set for the parameter probability of breaking the relationship, the variable engaged? is set again false for both agents.

- *Get influenced:*

Whether the agent is engaged or not it will be subject of social influence. If it is engaged its trait will become equal to the value of the trait of its partner, weighted by the value of parameter weight, plus the median of the values of their local relationships, weighted by one minus weight. When it is not engaged, the value of its trait just becomes the median of the value of their local relationships.

REFERENCES

[1] J. S. Coleman, E. Katz and H. Menzel., "The Diffusion of an Innovation among Physicians", *Sociometry* vol. 20, pp. 253-270, 1957.

[2] M. Rolfe, "Conditional Choice" in *The Oxford Handbook of Analytical Sociology*, P. Hedström and P. Bearman, Eds. Oxford: Oxford University Press, 2009, pp. 419 - 446.

[3] M. McPherson, L. Smith-Lovin and J. M. Cook, "Birds of a Feather: Homophily in Social Networks", *Annual Review of Sociology*, Vol. 27, pp. 415 – 44, 2001.

[4] Bearman, Peter, J. Moody and K. Stovel. 2004. "Chains of Affection. The Structure of Adolescent Romantic and Sexual Netwroks". *American Journal of Sociology*, 110: 44-91.

[5] Blau, Peter M. 1977. "A Macrosociological Theory of Social Structure", American Journal of Sociology, 83: 26-54.

[6] Feld, S. L. 1981. "The Focused Organization of Social Ties". American Journal of Sociology 86 (5): 1015-1035.

[7] Kossinets, G & D. J. Watts. 2009. "Origins of Homophily in an Evolving Social Network", American Journal of Sociology 115 (2): 405-50.

[8] S. F. Railsback and V. Grimm, *Agent-based and Individual-based Modeling*. Princeton, NY: Princeton University Press, 2012.

[9] U. Wilenski, NetLogo. http://ccl.northwestern.edu/netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. 1998

[10] Schelling, Thomas. 1978. Micromotives and macrobehavior. New York: Norton.

[11] Axelrod, Robert. 1997. "The dissemination of culture: a model with local convergence and global polarization". Journal of Conflict Resolution, 41 (2): 203-226.